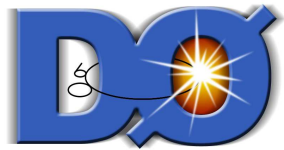


Remote Computing



Daniel Wicke
(Bergische Universität Wuppertal)

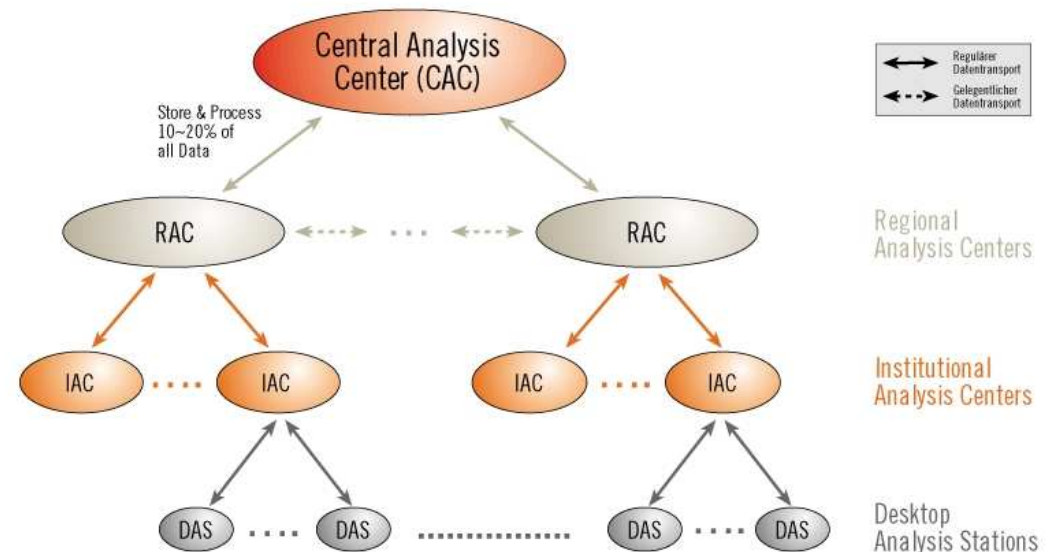


Outline

- Introduction
- Data Reprocessing
- Monte Carlo Production
- Summary

Introduction

- Remote computing has been in DØ's plan since ~ 1997 .
 - All Monte Carlo for RunII has been produced remotely.
 - SAM to be used for data handling.
- Since 2002 DØ is increasing its offsite computing usage:
 - Regional Analysis Centers established a tiered structure for data access.
 - Allows (manual) remote production and analysis
- Increasing use of the GRID
 - Monte Carlo
 - Data Processingwith unified/centralised submission.



Data Reprocessing

Improved detector understanding and new algorithms require rereconstruction

Computing Task

	2005 (p17)	2003/4 (p14)
Luminosity	470 pb ⁻¹	100 pb ⁻¹
Events	1G	300M
Rawdata 250kB/Event	250TB	75TB
DSTs 150kB/Event	150TB	45TB
TMBs 70(20)kB/Event	70TB	6TB
Time 50s/Event	20,000months	6000months
(on 1GHz Pentium III)	3400CPUs for 6mths	2000CPUs for 3mths
Remote processing	100%	30%

Central Farm (1000CPUs) used to capacity with data taking.

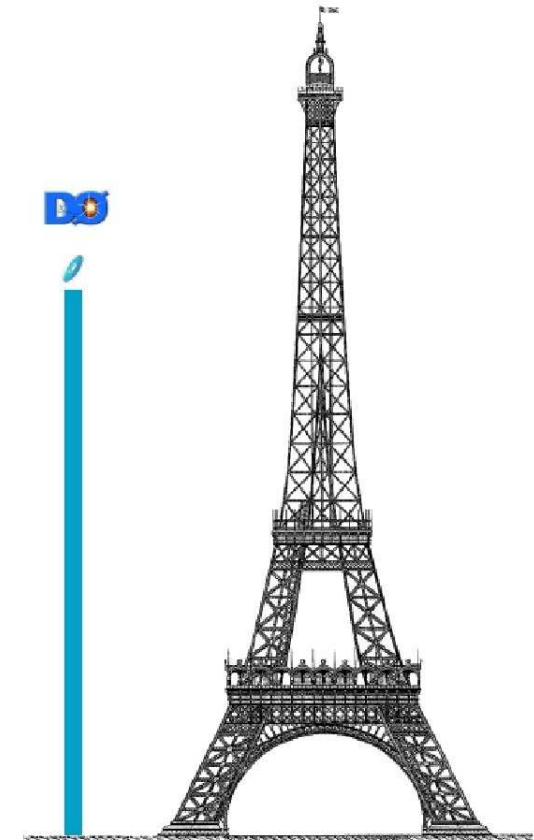
Data Reprocessing

Improved detector understanding and new algorithms require rereconstruction

Computing Task

Luminosity	2005 (p17)
Events	470 pb ⁻¹
Rawdata 250kB/Event	1G
DSTs 150kB/Event	250TB
TMBs 70(20)kB/Event	150TB
Time 50s/Event	70TB
(on 1GHz Pentium III)	20,000months
Remote processing	3400CPUs for 6mths
	100%

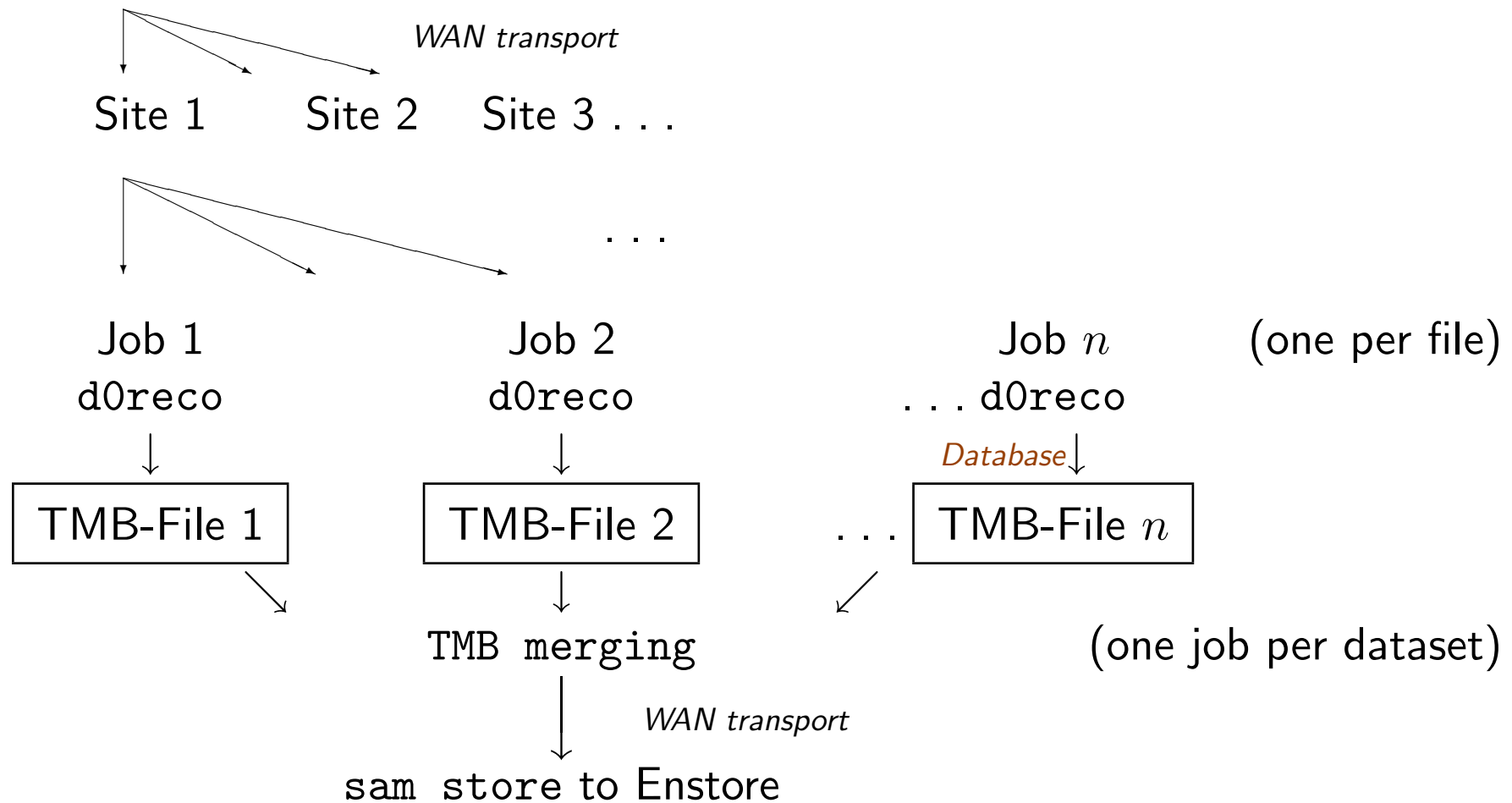
A stack of CDs as high as the Eiffeltower



Application flow

Overview

Datasets of RAW-files



Implementation

SAMGrid was chosen to implement this task on distributed systems.

- provides common environment for d0reco at all sites.
- allows common operation scripts (d0repro).

Production Step

- Each dataset processed through d0reco in one grid job.
- The grid jobs spawns one batch job per input file
- Resulting intermediate files are stored to SAM durable location (disk)

Scalability was improved by a factor of 500 to 1000(!)

Merging Step

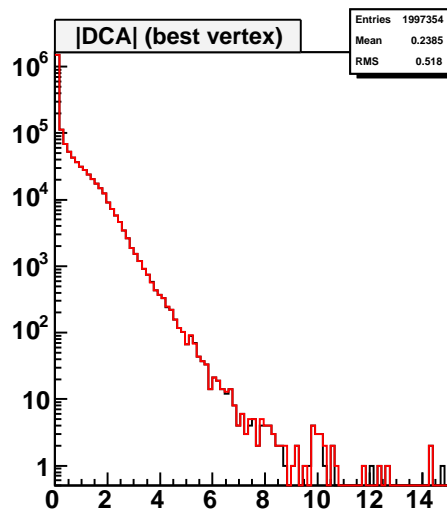
- Merge TMBs after all RAW-files of a run, $\mathcal{O}(100)$, are successfully processed.
 - But there are crashed and failures.
- ⇒ Merge only those that succeed; recover independently.
- Book-keeping is essential to avoid merging one TMB into two merged-TMBs.

At any stage SAM will know what happened to a file

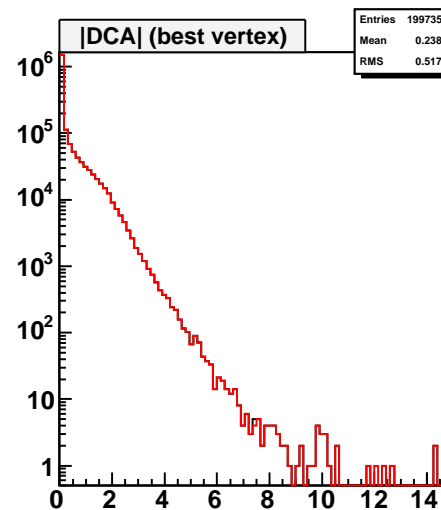
Certification of Sites and Code

- Compared SAMGrid production to conventional production on d0farm.
- Compared SAMGrid production at each site to d0farm production.
- Compared merged to unmerged TMBs at each site.

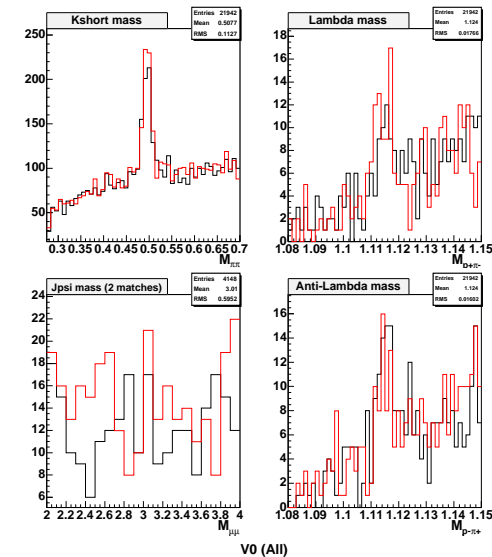
Lead to significant improvements in recocert



D0Farm JIM vs Lyon JIM



D0Farm Std vs JIM



Certified Resources

DØFarm (Fermilab),	1000CPUs	
Lyon	400CPUs	
Westgrid (Vancouver),	600CPUs	
Wisconsin	30CPUs	not in DØ
Prague	200CPUs	
DØSAR (UTA, Osker, Sprace)	510CPUs	(Osker: shared with meteorology and others)
CMS Farm (Fermilab)	300CPUs	OSG, JIM, SAM external; not DØ
UK (London, Man., Lanc., RAL)	750CPUs	only 250 certified
GridKa	500CPUs	recurring local problems
External	~3290CPUs	(1GHz PIII equiv.)

⇒ *not* sufficient for completion in 6mths,

but dØreco faster than projected;

opportunistic resources could be used at especially at Lyon and Westgrid.

Operational Tools

The application flow starts with a dataset to be processed.

Requirements

Deal with operations **including procedures for all possible failure modes.**

- submit full or partial set of production jobs.
- submit the corresponding (partial) merging job.
- determine the full or partial success of either step.
- (on request) create and submit the recovery jobs for both steps in case of (partial) failure.

Implemented using

- SAM for obtaining the information about files and
- JIM to submit jobs.

These scripts are common to all sites

D0Repro (Basic commands)

- Support for certification
- Submission (and recovery) is done by
`sub_production.py <dataset> <d0release>`
`sub_merge.py <dataset> <d0release>`
- Determination of production and merge status (poor man's request system)
`check_production.py <dataset> <d0release>`
`check_merge.py <dataset> <d0release>`
- Manually modify status of jobs
`set_status.py [production|merge] [approved|held|finished] <dataset> ...`

Typical workflow:

- 1) `sub_production.py ...` (investigate/retry in case of failures)
- 2) `sub_merge.py ...` (after production is finished; retry if failed)
- 3) `set_status.py ... finished ...` (in case of unrecoverable failures)

D0Repro (Autopilot functionalities)

- Investigate status of all active requests `check_all.py`
- Clean completed/finished datasets `clean_completed.py`
- Display status of all active requests and suggests `auto_pilot.py`
 - recover production if less than 5% failed
 - submit merge if unmerged files exist and last job was production
 - optionally approved additional production jobs (one per automatic merge submission)
- Run commands suggested by autopilot `source Autopilot.sh`

This chain could be run in a loop (with 1 or 2 hours delay).

Autopilot was built on the experience of reprocessing.

Significantly reduced work-load of operations

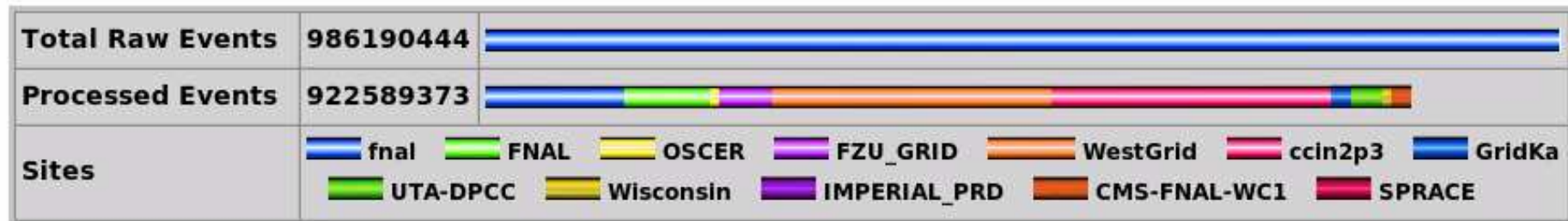
More than 90% of the operational work is to chase and fix failures.

Reliable book-keeping (taken from SAM) is prerequisite to implement these tools.

Status

Reprocessing effort started on 25-March-2005 in Lyon and Westgrid.

P17 Reprocessing Status as of 18-Oct-2005 (all sites)

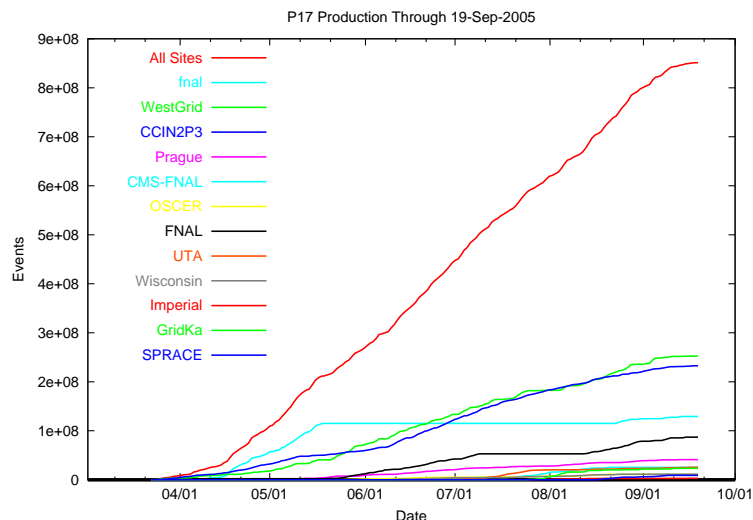


P17 Reprocessing Status as of 18-Oct-2005 (Remote sites only)

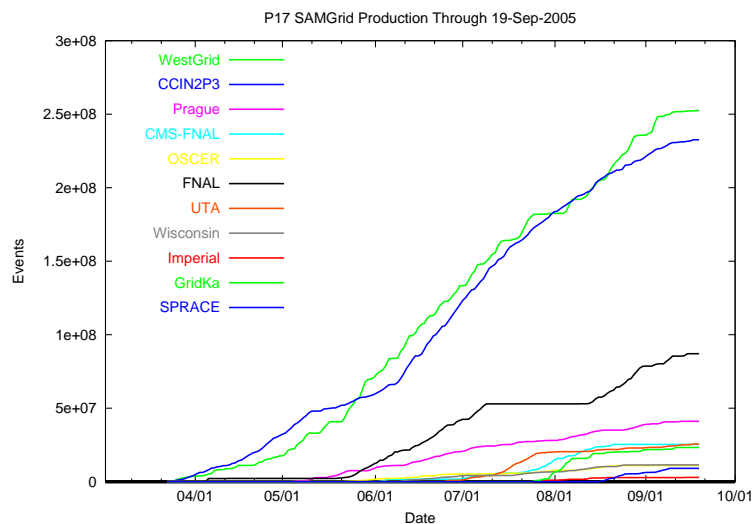


As of 18th October 923M of 986M events are completed, i.e. 94% done.

Integrated number of events (from SAM)



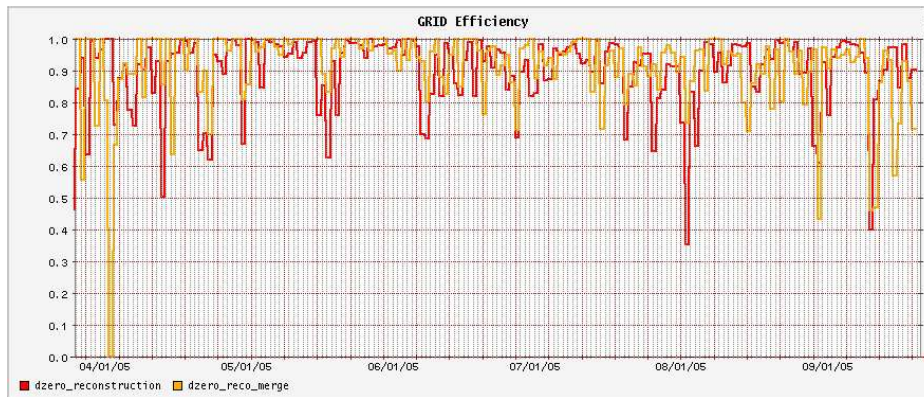
- 923M events completed
- Deployment of improved infrastructure visible as kink (\sim 25th Apr)
- Started at \sim 2.5MEvts/day.
- Reached up to \sim 10MEvts/day.



- 690M (75%) done remotely.
- 230M at FNAL (DØ and CMS farms).
- 790M with Grid Mechanisms..

Bulk production completed.

Efficiency



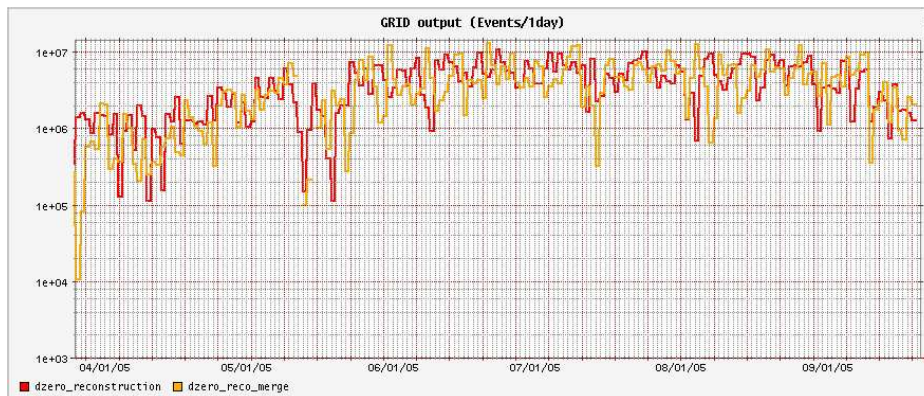
Efficiency of batch jobs that produced a file over the number of started jobs.

Average job failure rate 10%(!)*

Dominated by failures of services:
(Broken SAM, partial broken nodes, ...)

Rate of unrecoverable failures 0.2%**

Speed



Production Speed in Events/day*

Needed 6M/day for 1G in 6 months.

*Based on XML (by construction pessimistic). ** based on SAM

Monte Carlo Production

- Over the last year DØ produced around 80M Monte Carlo events ($\sim 7\text{TB}$).
- These were produced at 10 different remote sites:

Resources

	Submission Mechanism
DØSAR (LTU, Luhep, Osker, Ouhep, Sprace, Tata, UTA)	SAMGrid
FZU (Prague)	SAMGrid
GridKa	SAMGrid
IN2P3, Lyon	local/SAMGrid
Nikhef	local/LCG
Wisconsin	SAMGrid
DØFarm	SAMGrid
UK (Imperial, Manchester)	SAMGrid
Westgrid	SAMGrid

(Most of the resources are shared)

- SamGrid in use since spring 2004
- Developement was focused on reprocessing
- Now emphasis back to MC
- Reprocessing resources can be used without reconfiguration

Summary

- Data reprocessing effort more than $3\times$ bigger than the 2003/4 effort.
 - 250TB; 1600CPU years. Largest distributed HEP effort.
 - Fully gridified, common tools, 11 sites.
 - On schedule (i.e. bulk prod. done).
- Monte Carlo Production
 - moving from distributed computing to a gridified operation.
 - exploiting opportunistic resources
OSG, LCG, d0farm
- Grid is starting to return some investment
 - person power intense setup
 - common submission tools
 - sites installed for reprocessing
can be used for MC

